

# On Wu and Schaback's Error Bound

Lin-Tian Luh

February 2, 2008

AMS classifications: 41A05, 41A63, 41A25, 65D05

Keywords: multivariate interpolation, Kriging function, multiquadrics, Gaussians, thin-plate splines.

## Abstract

Radial basis functions are a very powerful tool in multivariate approximation since they are mesh free. In the theory of radial basis functions, the most frequently used error bounds are the one raised by Madych and Nelson in [5] and the one raised by Wu and Schaback in [6], especially the latter. It seems that Wu and Schaback's error bound is five times more frequently used than that of Madych and Nelson's. The reason is that Wu and Schaback's space of approximated functions is much easier to understand, and their error bound is much easier to use. Unfortunately, this error bound contains crucial mistakes which will be pointed out in this paper. Moreover, the r.b.f. people have misunderstood its space of approximands for long. This will also be discussed in depth. These, to some extent, may be a blow to r.b.f. people. However, Madych and Nelson's error bound is still very powerful because it's highly related to Sobolev spaces which contain solutions of many differential equations. All these discussions are based on the central idea of this paper, i.e. unsmooth functions should be approximated by unsmooth functions.

## 1 Introduction

Radial basis function interpolation to scattered data  $(x_i, f_i) \in R^{n+1}$  for pairwise distinct points ("centers")  $x_1, \dots, x_M \in R^n$  uses a function  $\phi : R_{\geq 0} \rightarrow R$  and the space  $P_q$  of polynomials on  $R^n$  with order not exceeding  $q$  to construct the interpolant

$$s(x) := \sum_{i=1}^M a_i \phi(\|x - x_i\|) + \sum_{i=1}^Q b_i p_i(x) \quad (1)$$

via the linear system

$$\begin{aligned} \sum_{i=1}^M a_i \phi(\|x_j - x_i\|) + \sum_{i=1}^Q b_i p_i(x_j) &= f_j, 1 \leq j \leq M \\ \sum_{j=1}^M a_j p_i(x_j) &= 0, \quad 1 \leq i \leq Q = \binom{q+n-1}{n} \end{aligned} \quad (2)$$

where  $p_1, \dots, p_Q$  are a basis of  $P_q$ , and  $f_j$  means  $f(x_j)$  with  $f$  being the approximated function. For a wide choice of  $\phi$  and polynomial orders  $q$ , including the case  $q = Q = 0$ , the nonsingularity of the  $(M+Q) \times (M+Q)$  system (2), written as

$$\begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \quad (3)$$

in matrix notation, has been established by r.b.f. people. A sufficient condition for the linear system to be solvable is that  $\psi(x) := \phi(\|x\|)$  should be conditionally positive definite of order  $q$  (see [2] for its definition).

The approximated function  $f$  should be defined on  $\Omega \subseteq R^n$ . There are no further conditions on  $\Omega$ . The sampling points  $x_i$  are allowed to be irregularly distributed over  $\Omega$  under the restriction

$$p(x_j) = 0, 1 \leq j \leq M \text{ for } p \in P_q \text{ implies } p = 0 \quad (4)$$

, which guarantees the uniqueness of the solution of the linear system. For local error estimation, we measure the “density” of centers  $x_j$  near some  $x \in \Omega$  by

$$h_\rho(x) := \max_{y \in K_\rho(x)} \min_{1 \leq j \leq M} \|y - x_j\|$$

for some fixed  $\rho > 0$ , where  $K_\rho(x) = \{y \in R^n : \|x - y\| \leq \rho\}$ , using the Euclidean norm  $\|\cdot\|$ .

## 2 A brief description of Wu and Schaback’s results

Wu and Schaback showed local error bounds of the following form: Given constants  $\rho \in R_{>0}$ ,  $q \in N_{\geq 0}$ , a radial basis function  $\phi$  and a certain function space  $F_\phi$  to be described below, there exist positive constants  $k, m \in N$ ,  $k \geq m$ ,  $h_0, C \in R$  such that for any distribution of points  $x_j \in R^n$ ,  $1 \leq j \leq M$ , any function  $f$  from  $F_\phi$  and any point  $x \in \Omega$  with  $h_\rho(x) < h_0$ , the inequality

$$|s^{(\mu)}(x) - f^{(\mu)}(x)| \leq c_f \cdot C \cdot h_\rho^{k-|\mu|}(x) \quad (5)$$

holds for the error and its  $\mu$ -th derivatives for  $0 \leq |\mu| \leq m$ , where the constant  $c_f$  depends on  $f, \phi$ , and  $F_\phi$  only. Then (5) is called a **local error bound** of order  $k$ . We define  $|\mu| := \sum_i \mu_i$  for  $\mu \in N^n$ .

The linear equations (2) are solvable under the hypotheses of the preceding section. It follows that there is a Lagrange-type representation

$$s(x) = \sum_{j=1}^M f(x_j) u_j(x), \quad u_j(x_i) = \delta_{ij}, \quad 1 \leq i, j \leq M, \quad (6)$$

of the solution. Introducing vectors

$$R(x) := (\phi(\|x - x_1\|), \dots, \phi(\|x - x_M\|))^T \quad (7)$$

$$S(x) := (p_1(x), \dots, p_Q(x))^T, \quad (8)$$

we get the following theorem.

**Theorem 2.1** *The vector  $U(x) := (u_1(x), \dots, u_M(x))^T$  formed by the values of the Lagrange basis functions  $u_1, \dots, u_M$  of (6) at  $x \in R^n$  coincides with the solution  $U_*(x)$  of the conditional minimization problem*

$$\min\{U^T A U - 2U^T R(x) + \phi(0) \mid U \in R^M, P^T U = S(x)\}. \quad (9)$$

Proof: omitted.

We can do the same thing for derivatives. If  $\phi$  is differentiable of order  $|\mu|$  on  $(0, \infty)$  and of order  $2|\mu|$  around zero, and if  $\psi(x) := \phi(\|x\|)$ , the solution  $U_*^{(\mu)}(x)$  of the problem

$$\min\{U^T A U - 2U^T R^{(\mu)}(x) + \psi^{(2\mu)}(0) \mid U \in R^M, P^T U = S^{(\mu)}(x)\} \quad (10)$$

uniquely exists. Moreover, the derivative  $U^{(\mu)}(x)$  of  $U(x)$  coincides with the solution  $U_*^{(\mu)}(x)$  of (10). It's worth noting that

$$\sum_{j=1}^M u_j^{(\mu)}(x) \cdot p(x_j) = p^{(\mu)}(x) \quad (11)$$

for all  $p \in P_q$ . The choice of the additive constant in (10) is required for the following definition.

**Definition 2.1** *Let  $\psi(x) = \phi(\|x\|)$  be conditionally positive definite of order  $q$ , and assume  $\phi \in C^{|\mu|}(0, \infty)$ ,  $\phi \in C^{2|\mu|}$  around zero for  $\mu \in N_{\geq 0}^n$ . Then, for any distribution of centers  $x_1, \dots, x_M$  satisfying (4), the nonnegative function  $\kappa_q^{(\mu)}$  defined by*

$$(\kappa_q^{(\mu)}(x))^2 := \min\{U^T A U - 2U^T R^{(\mu)}(x) + \psi^{(2\mu)}(0) \mid U \in K_q^{(\mu)}(x)\}$$

with the set

$$K_q^{(\mu)} := \{U = (u_1, \dots, u_M)^T \in R^M \mid \sum_{j=1}^M u_j p(x_j) = p^{(\mu)}(x) \text{ for all } p \in P_q\} \quad (12)$$

of admissible vectors is called the **Kriging function** at  $x$ .

The following theorem is the prototype of Wu and Schaback's main result.

**Theorem 2.2** *If the hypotheses of Definition 2.1 hold and the Fourier transform  $\hat{\psi}$  of  $\psi$  satisfies*

$$\psi(y) = \phi(\|y\|) = \frac{1}{(2\pi)^n} \int_{R^n} e^{i\langle y, t \rangle} \hat{\psi}(t) dt, \quad y \in R^n, \quad (13)$$

then

$$U^T A U - 2U^T R^{(\mu)}(x) + \psi^{(2\mu)}(0) = \frac{1}{(2\pi)^n} \int_{R^n} \left| \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle} \right|^2 \hat{\psi}(t) dt \quad (14)$$

for arbitrary  $U = (u_1, \dots, u_M) \in R^M$ , expressing the Kriging function via an integral.

Proof: The theorem follows from

$$\begin{aligned} & \sum_{j,k=1}^M w_j \overline{w_k} \phi(\|x_j - x_k\|) \\ &= \frac{1}{(2\pi)^n} \int_{R^n} \sum_{j,k=1}^M w_j \overline{w_k} e^{i\langle x_j - x_k, t \rangle} \hat{\psi}(t) dt \\ &= \frac{1}{(2\pi)^n} \int_{R^n} \left| \sum_{j=1}^M w_j e^{i\langle x_j, t \rangle} \right|^2 \hat{\psi}(t) dt \end{aligned}$$

and

$$\begin{aligned} \frac{d^\mu}{dx^\mu} \phi(\|x - x_j\|) &= \frac{d^\mu}{dx^\mu} \frac{1}{(2\pi)^n} \int_{R^n} e^{i\langle x - x_j, t \rangle} \hat{\psi}(t) dt \\ &= \frac{1}{(2\pi)^n} \int_{R^n} (it)^\mu e^{i\langle x - x_j, t \rangle} \hat{\psi}(t) dt \end{aligned}$$

△

However, for most of the interesting radial basis functions  $\phi$ , (13) does not hold. Hence we must develop an approach of generalized functions.

**Theorem 2.3** *Let the generalized Fourier transform of  $\psi(x) = \phi(\|x\|)$  exist and coincide with a continuous function  $\hat{\psi}$  on  $R^n - \{0\}$  satisfying*

$$0 < \hat{\psi}(t) \leq c \begin{cases} \|t\|^{-n-s_0} & \text{for } \|t\| \longrightarrow 0 \\ \|t\|^{-n-s_\infty} & \text{for } \|t\| \longrightarrow \infty \end{cases} \quad (15)$$

with constants  $c \in R_{>0}$ ,  $s_0, s_\infty \in R$ , where we additionally assume

$$2|\mu| < s_\infty \text{ and } s_0 < 2q \quad (16)$$

Then for all  $U \in K_q^{(\mu)}(x)$  we have (14) with a well-defined integral.

Proof: Let  $g_U(t) := \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle}$ . By the theorem on monotone convergence,

$$\begin{aligned} \frac{1}{(2\pi)^n} \int_{R^n} |g_U(t)|^2 \cdot |\hat{\psi}(t)| dt &= \lim_{m \rightarrow \infty} \frac{1}{(2\pi)^n} \int_{R^n} |g_U(t)|^2 e^{\frac{-\|t\|^2}{m^2}} \hat{\psi}(t) dt \\ &= \lim_{m \rightarrow \infty} \frac{1}{(2\pi)^n} \int_{R^n} \hat{G}_m(z) dz \end{aligned} \quad (17)$$

with the test functions

$$G_m(t) = |g_U(t)|^2 e^{\frac{-\|t\|^2}{m^2}}$$

using the definition of the generalized Fourier transform for tempered distributions. Furthermore, the Fourier transform  $\hat{G}_m(z)$  of  $G_m$  can be explicitly calculated up to a constant  $\sigma_{mn}$  as

$$\begin{aligned} &m^n \sum_{j,k} u_j u_k e^{\frac{-\|z - (x_j - x_k)\|^2 m^2}{4}} \\ &- 2m^n (-1)^{|\mu|} \sum_k u_k D^\mu \left( e^{\frac{-\|z - (x - x_k)\|^2 m^2}{4}} \right) \\ &+ m^n D^{2\mu} \left( e^{\frac{-\|z\|^2 m^2}{4}} \right), \end{aligned}$$

where  $D$  denotes differentiation with respect to  $z$ . Insertion into (17) yields (14).  $\triangle$

We can now write the Kriging function as an integral:

**Theorem 2.4** *Under the assumptions of Theorem 2.3 the Kriging function has the representation*

$$(\kappa_q^{(\mu)}(x))^2 = \min_{U \in K_q^{(\mu)}(x)} \frac{1}{(2\pi)^n} \int_{R^n} \left| \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle} \right|^2 \hat{\psi}(t) dt \quad (18)$$

with  $K_q^{(\mu)}$  defined as in (12), and the integral exists in the classical sense.  $\triangle$

**Lemma 2.1** *If the data  $(x_j, f_j)$  stem from an absolutely integrable real-valued function  $f$  on  $R^n$  with a nicely behaving Fourier transform  $\hat{f}$  satisfying*

$$f(x) = \frac{1}{(2\pi)^n} \int_{R^n} e^{i\langle x, t \rangle} \hat{f}(t) dt, \quad x \in R^n, \quad (19)$$

then

$$|s^{(\mu)}(x) - f^{(\mu)}(x)|^2 = \left| \frac{1}{(2\pi)^n} \int_{R^n} \left( \sum_{j=1}^M u_j^{(\mu)}(x) e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle} \right) \hat{f}(t) dt \right|^2 \quad (20)$$

where  $\phi, \mu$  are as in Definition 2.1 and  $u_1(x), \dots, u_M(x)$  are Lagrange interpolation functions.

Proof: Note that

$$\begin{aligned} s^{(\mu)}(x) - f^{(\mu)}(x) &= \sum_{j=1}^M f(x_j) u_j^{(\mu)}(x) - f^{(\mu)}(x) \\ &= \sum_{j=1}^M u_j^{(\mu)}(x) \frac{1}{(2\pi)^n} \int_{R^n} e^{i\langle x_j, t \rangle} \cdot \hat{f}(t) dt - \frac{1}{(2\pi)^n} \int_{R^n} (it)^\mu e^{i\langle x, t \rangle} \hat{f}(t) dt \\ &= \frac{1}{(2\pi)^n} \int_{R^n} \left( \sum_{j=1}^M u_j^{(\mu)}(x) e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle} \right) \hat{f}(t) dt. \end{aligned}$$

The theorem follows immediately.  $\triangle$

Assume that

$$c_f^2 := \frac{1}{(2\pi)^n} \int_{R^n} |\hat{f}(t)|^2 (\hat{\psi}(t))^{-1} dt < \infty \quad (21)$$

in addition to the hypotheses of Definition 2.1 and Theorem 2.3. Then using the Cauchy-Schwarz inequality, we get

$$\begin{aligned} &|s^{(\mu)}(x) - f^{(\mu)}(x)|^2 \\ &= \frac{1}{(2\pi)^n} \int_{R^n} \left| \sum_{j=1}^M u_j^{(\mu)}(x) e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle} \right|^2 \hat{\psi}(t) dt \cdot \frac{1}{(2\pi)^n} \int_{R^n} |\hat{f}(t)|^2 (\hat{\psi}(t))^{-1} dt \\ &= (\kappa_q^{(\mu)}(x))^2 \cdot c_f^2. \end{aligned} \quad (22)$$

To cope with generalized Fourier transforms, we proceed as before and need the following definition.

**Definition 2.2** A function  $f : R^n \rightarrow R$  is dominated by a radial basis function  $\phi$  satisfying (15) and (16) on  $R^n - \{0\}$ , iff  $f$  has a generalized Fourier transform  $\hat{f}$  coinciding on  $R^n - \{0\}$  with a continuous function satisfying (21) for  $\psi(x) = \phi(\|x\|)$ .

Remark: The set  $F_\phi$  of functions dominated by  $\phi$  may be completed to form a Hilbert space with inner product

$$(f_1, f_2)_\phi = \int_{R^n} \hat{f}_1(t) \overline{\hat{f}_2(t)} (\hat{\psi}(t))^{-1} dt$$

which was thoroughly studied by Madych and Nelson in [4], [5].

**Theorem 2.5** *If  $f$  is in the space  $F_\phi$  of functions dominated by a radial basis function  $\phi$  satisfying (15) and (16), then the interpolation error can be bounded by*

$$|s^{(\mu)}(x) - f^{(\mu)}(x)| \leq \kappa_q^{(\mu)}(x) \cdot c_f \quad (23)$$

, where  $c_f$  is given by (21).

Proof: This is an immediate result of the above arguments.  $\triangle$

**Theorem 2.6** *Let  $\phi$  satisfy the assumptions of Theorem 2.3, and let  $\rho \in R_{>0}$  be given. Then there exist positive real constants  $h_0$  and  $C$  such that for any distribution of centers  $x_i \in R^n$ ,  $1 \leq i \leq M$ , and any point  $x \in R^n$  with  $h_\rho(x) \leq h_0$  the Kriging function can be bounded by*

$$\kappa_q^{(\mu)}(x) \leq C \cdot h_\rho^{\frac{s_\infty}{2} - |\mu|}(x) \quad (24)$$

for  $0 \leq |\mu| \leq \frac{s_\infty}{2}$ .

Proof: omitted.  $\triangle$

### 3 Crucial mistakes

The first mistake occurs in (14). The formula holds only when  $|\mu|$  is even. Let's check the equation in detail.

$$\begin{aligned} & U^T A U - 2U^T R^{(\mu)}(x) + \psi^{(2\mu)}(0) \\ &= \frac{1}{(2\pi)^n} \int_{R^n} \left| \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} \right|^2 \hat{\psi}(t) dt - 2 \sum_{j=1}^M u_j \cdot \frac{1}{(2\pi)^n} \int_{R^n} (it)^\mu e^{i\langle x - x_j, t \rangle} \hat{\psi}(t) dt \\ & \quad + \psi^{(2\mu)}(0) \\ &= \frac{1}{(2\pi)^n} \int_{R^n} \left| \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} \right|^2 \hat{\psi}(t) dt - \frac{1}{(2\pi)^n} \int_{R^n} 2 \sum_{j=1}^M u_j (it)^\mu e^{i\langle x - x_j, t \rangle} \hat{\psi}(t) dt \\ & \quad + \psi^{(2\mu)}(0). \end{aligned}$$

Now,

$$\begin{aligned}
& \frac{1}{(2\pi)^n} \int_{R^n} \left| \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle} \right|^2 \hat{\psi}(t) dt \\
&= \frac{1}{(2\pi)^n} \int_{R^n} \left[ \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle} \right] \left[ \sum_{j=1}^M \overline{u_j} e^{-i\langle x_j, t \rangle} - \overline{(it)^\mu} e^{-i\langle x, t \rangle} \right] \hat{\psi}(t) dt \\
&= \frac{1}{(2\pi)^n} \int_{R^n} \left\{ \left| \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} \right|^2 - \sum_{j=1}^M u_j \overline{(it)^\mu} e^{i\langle x_j - x, t \rangle} - \sum_{j=1}^M \overline{u_j} (it)^\mu e^{i\langle x - x_j, t \rangle} \right. \\
&\quad \left. + (it)^\mu \cdot \overline{(it)^\mu} \right\} \hat{\psi}(t) dt \\
&= \frac{1}{(2\pi)^n} \int_{R^n} \left| \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} \right|^2 \hat{\psi}(t) dt - \frac{1}{(2\pi)^n} \int_{R^n} \sum_{j=1}^M u_j \overline{(it)^\mu} e^{i\langle x_j - x, t \rangle} \hat{\psi}(t) dt \\
&\quad - \frac{1}{(2\pi)^n} \int_{R^n} \sum_{j=1}^M \overline{u_j} (it)^\mu e^{i\langle x - x_j, t \rangle} \hat{\psi}(t) dt + (-1)^{|\mu|} \psi^{(2\mu)}(0) \\
&= \frac{1}{(2\pi)^n} \int_{R^n} \left| \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} \right|^2 \hat{\psi}(t) dt - \frac{1}{(2\pi)^n} \int_{R^n} \sum_{j=1}^M u_j \overline{(it)^\mu} e^{i\langle x_j - x, -t \rangle} \hat{\psi}(-t) dt \\
&\quad - \frac{1}{(2\pi)^n} \int_{R^n} \sum_{j=1}^M \overline{u_j} (it)^\mu e^{i\langle x - x_j, t \rangle} \hat{\psi}(t) dt + (-1)^{|\mu|} \psi^{(2\mu)}(0) \\
&= \begin{cases} U^T A U - 2U^T R^{(\mu)}(x) + (-1)^{|\mu|} \psi^{(2\mu)}(0) & \text{whenever } |\mu| \text{ is even,} \\ U^T A U + (-1)^{|\mu|} \psi^{(2\mu)}(0) & \text{whenever } |\mu| \text{ is odd.} \end{cases}
\end{aligned}$$

Note that it's better to add a constant  $(-1)^{|\mu|}$  to  $\psi^{(2\mu)}(0)$  in the definition of the Kriging function.

The second mistake occurs in the proof of Theorem 2.3. There are three problems in the proof. The first problem is that the Fourier transform  $\hat{G}_m(z)$  of  $G_m$  is wrong. Let's show it.

$$\begin{aligned}
& \hat{G}_m(z) \\
&= \frac{1}{(2\pi)^n} \int_{R^n} G_m(t) e^{-i\langle z, t \rangle} dt \\
&= \frac{1}{(2\pi)^n} \int_{R^n} |g_U(t)|^2 e^{-\|t\|^2/m^2} \cdot e^{-i\langle z, t \rangle} dt \\
&= \frac{1}{(2\pi)^n} \int_{R^n} \left[ \sum_{j=1}^M u_j e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle} \right] \left[ \sum_{j=1}^M \overline{u_j} e^{-i\langle x_j, t \rangle} - \overline{(it)^\mu} e^{-i\langle x, t \rangle} \right] \\
&\quad e^{-\|t\|^2/m^2} \cdot e^{-i\langle z, t \rangle} dt
\end{aligned}$$



$$\begin{aligned}
&= \frac{1}{(2\pi)^n} \int_{R^n} \left[ \sum_{j,k} u_j u_k e^{i\langle x_j - x_k, t \rangle} + \sum_{j=1}^M u_j \overline{(it)^\mu} e^{i\langle x_j - x, t \rangle} \right. \\
&\quad \left. - \sum_{j=1}^M \overline{u_j} (it)^\mu e^{i\langle x - x_j, t \rangle} + (it)^\mu \overline{(it)^\mu} \right] e^{-\|t\|^2/m^2} \cdot e^{-i\langle z, t \rangle} dt \\
&= \frac{1}{(2\pi)^n} \int_{R^n} \sum_{j,k} u_j u_k e^{-\|t\|^2/m^2} \cdot e^{-i\langle z - (x_j - x_k), t \rangle} dt \\
&\quad + \frac{1}{(2\pi)^n} \int_{R^n} \sum_{j=1}^M u_j \overline{(it)^\mu} e^{-\|t\|^2/m^2} \cdot e^{-i\langle z - (x_j - x), t \rangle} dt \\
&\quad - \frac{1}{(2\pi)^n} \int_{R^n} \sum_{j=1}^M \overline{u_j} (it)^\mu \cdot e^{-\|t\|^2/m^2} \cdot e^{-i\langle z - (x - x_j), t \rangle} dt \\
&\quad + \frac{1}{(2\pi)^n} \int_{R^n} (it)^\mu \overline{(it)^\mu} e^{-\|t\|^2/m^2} \cdot e^{-i\langle z, t \rangle} dt \\
&= \sigma_{mn} m^n \sum_{j,k} u_j u_k e^{-\|z - (x_j - x_k)\|^2 m^2/4} \\
&\quad + \sigma_{mn} m^n \sum_{j=1}^M u_j D^\mu e^{-\|z - (x_j - x)\|^2 m^2/4} \\
&\quad + \sigma_{mn} m^n (-1)^{|\mu|} \sum_{j=1}^M u_j D^\mu e^{-\|z - (x - x_j)\|^2 m^2/4} \\
&\quad + \sigma_{mn} m^n (-1)^{|\mu|} D^{2\mu} e^{-\|z\|^2 m^2/4}
\end{aligned}$$

, where  $\sigma_{mn}$  is some constant independent of  $m$ .

Note that the second and third items of the result cannot be incorporated whenever  $|\mu|$  is odd. Now, the second problem arises when we insert  $\hat{G}_m$  into (17). We display the process.

$$\begin{aligned}
&\lim_{m \rightarrow \infty} \frac{1}{(2\pi)^n} \int_{R^n} \hat{G}_m(z) \psi(z) dz \\
&= \lim_{m \rightarrow \infty} \frac{1}{(2\pi)^n} \cdot \sigma_{mn} \sum_{j,k} u_j u_k \int_{R^n} m^n e^{-\|z - (x_j - x_k)\|^2 m^2/4} \cdot \psi(z) dz \\
&\quad + \lim_{m \rightarrow \infty} \frac{1}{(2\pi)^n} \cdot \sigma_{mn} \sum_{j=1}^M \int_{R^n} m^n D^\mu e^{-\|z - (x_j - x)\|^2 m^2/4} \cdot \psi(z) dz \\
&\quad + \lim_{m \rightarrow \infty} \frac{1}{(2\pi)^n} \cdot \sigma_{mn} (-1)^{|\mu|} \sum_{j=1}^M \int_{R^n} m^n D^\mu e^{-\|z - (x - x_j)\|^2 m^2/4} \cdot \psi(z) dz \\
&\quad + \lim_{m \rightarrow \infty} \frac{1}{(2\pi)^n} \cdot \sigma_{mn} (-1)^{|\mu|} \int_{R^n} m^n D^{2\mu} e^{-\|z\|^2 m^2/4} \cdot \psi(z) dz \quad (25)
\end{aligned}$$

Wu and Schaback conclude that the limits of the four integrals of the right-hand side are equal to

$$\psi(x_j - x_k), (-1)^{|\mu|} D^\mu \psi(x_j - x), (-1)^{|\mu|} D^\mu \psi(x - x_j) \text{ and } (-1)^{2|\mu|} D^{2\mu} \psi(0)$$

respectively, by applying  $\delta$ -function. However, by p.217 of [1],  $\psi$  should be a good function when applying  $\delta$ -function. That is,  $\psi$  should be a member of the Schwartz space. Fortunately, as long as we require  $\psi$  to be in  $C^1$ , then  $\delta$ -function still works even if  $\psi$  is not a good function. This can be shown as follows.

$$\int_{-\infty}^{\infty} (m/\pi)^{n/2} e^{-mr^2} \psi(x) dx = \psi(0) + \int_{-\infty}^{\infty} (m/\pi)^{n/2} \{\psi(x) - \psi(0)\} e^{-mr^2} dx.$$

Since

$$\begin{aligned} |\psi(x) - \psi(0)| &= \left| \int_0^{x_1} \frac{\partial \psi(t_1, 0, \dots, 0)}{\partial t_1} dt_1 + \int_0^{x_2} \frac{\partial \psi(x_1, t_2, 0, \dots, 0)}{\partial t_2} dt_2 + \dots \right. \\ &\quad \left. + \int_0^{x_n} \frac{\partial \psi(x_1, x_2, \dots, t_n)}{\partial t_n} dt_n \right| \\ &\leq M(|x_1| + |x_2| + \dots + |x_n|) \end{aligned}$$

, the integral does not exceed  $Mn/m^{1/2}$  which tends to zero as  $m \rightarrow \infty$  and

$$\int_{-\infty}^{\infty} \delta(x) \psi(x) dx = \psi(0)$$

is justified.

What's making trouble is the derivative. This is the third problem. We know that for any generalized function  $g$  and good function  $\gamma$ ,

$$\int_{R^n} \{\partial_1^{q_1} \dots \partial_n^{q_n} g(x)\} \gamma(x) dx = (-1)^q \int_{R^n} g(x) \partial^q \gamma(x) dx$$

where  $q = (q_1, \dots, q_n)$ . The requirement that  $\gamma(x)$  be a good function is absolutely necessary and cannot be modified. Thus the second and third integrals in the right-hand side of (25) cannot be reduced to  $(-1)^{|\mu|} D^\mu \psi(x_j - x)$  and  $(-1)^{|\mu|} D^\mu \psi(x - x_j)$  respectively. Moreover, the item  $\psi^{(2\mu)}(0)$  needs a coefficient  $(-1)^{|\mu|}$  also. These bad mistakes destroy (18) thoroughly.

The third crucial mistake occurs in (22) whenever  $\hat{f}$  and  $\hat{\psi}$  are generalized functions. Note that (22) is a result of (20) which is also proven as follows for the generalized case.

Let  $g_U(t) = \sum_{j=1}^M u_j^{(\mu)}(x) e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle}$ . Then, by Lebesgue Convergence Theorem,

$$\frac{1}{(2\pi)^n} \int_{R^n} g_U(t) \hat{f}(t) dt = \lim_{m \rightarrow \infty} \frac{1}{(2\pi)^n} \int_{R^n} g_U(t) e^{-\|t\|^2/m^2} \hat{f}(t) dt$$

$$\begin{aligned}
&= \lim_{m \rightarrow \infty} \frac{1}{(2\pi)^n} \int_{R^n} \hat{G}_m(z) f(z) dz \\
&\text{where } G_m(t) := g_U(t) e^{-\|t\|^2/m^2}.
\end{aligned}$$

The Fourier transform  $\hat{G}_m(z)$  can be calculated as follows.

$$\begin{aligned}
\hat{G}_m(z) &= \frac{1}{(2\pi)^n} \int_{R^n} \left[ \sum_{j=1}^M u_j^{(\mu)}(x) e^{i\langle x_j, t \rangle} - (it)^\mu e^{i\langle x, t \rangle} \right] e^{-\|t\|^2/m^2} e^{-i\langle z, t \rangle} dt \\
&= \frac{1}{(2\pi)^n} \sum_{j=1}^M u_j^{(\mu)}(x) \int_{R^n} e^{-\|t\|^2/m^2} \cdot e^{-i\langle z-x_j, t \rangle} dt \\
&\quad - \frac{1}{(2\pi)^n} \int_{R^n} (it)^\mu e^{-\|t\|^2/m^2} \cdot e^{-i\langle z-x, t \rangle} dt \\
&= \frac{1}{(2\pi)^n} m^n \sum_{j=1}^M u_j^{(\mu)}(x) \cdot e^{-\frac{m^2}{4}\|z-x_j\|^2} - \frac{1}{(2\pi)^n} m^n (-1)^{|\mu|} D^\mu \int_{R^n} e^{-\frac{\|t\|^2}{m^2}} \cdot e^{-i\langle z-x, t \rangle} dt \\
&= \frac{1}{(2\pi)^n} m^n \sum_{j=1}^M u_j^{(\mu)}(x) \cdot e^{-\frac{m^2}{4}\|z-x_j\|^2} - \frac{1}{(2\pi)^n} m^n (-1)^{|\mu|} D_z^\mu e^{-\frac{m^2}{4}\|z-x\|^2}.
\end{aligned}$$

Now,

$$\begin{aligned}
&\lim_{m \rightarrow \infty} \frac{1}{(2\pi)^n} \int_{R^n} \hat{G}_m(z) f(z) dz \\
&= \lim_{m \rightarrow \infty} \sigma_{mn} [1/(2\pi)^n]^2 \int_{R^n} m^n \sum_{j=1}^M u_j^{(\mu)}(x) e^{-\frac{m^2}{4}\|z-x_j\|^2} \cdot f(z) - (-1)^{|\mu|} m^n D_z^\mu e^{-\frac{m^2}{4}\|z-x\|^2} \cdot f(z) dz \\
&= \frac{\sigma_{mn}}{(2\pi)^{2n}} \sum_{j=1}^M u_j^{(\mu)}(x) \lim_{m \rightarrow \infty} \int_{R^n} m^n e^{-\frac{m^2}{4}\|z-x_j\|^2} \cdot f(z) dz \\
&\quad - \frac{\sigma_{mn}}{(2\pi)^{2n}} \cdot (-1)^{|\mu|} \lim_{m \rightarrow \infty} \int_{R^n} m^n D_z^\mu e^{-\frac{m^2}{4}\|z-x\|^2} \cdot f(z) dz \\
&= \frac{\sigma_{mn}}{(2\pi)^{3n/2}} \sum_{j=1}^M u_j^{(\mu)}(x) f(x_j) - \frac{\sigma_{mn}}{(2\pi)^{3n/2}} (-1)^{|\mu|} \cdot (-1)^{|\mu|} \cdot D^\mu f(x) \\
&= \frac{\sigma_{mn}}{(2\pi)^{3n/2}} [s^{(\mu)}(x) - f^{(\mu)}(x)].
\end{aligned}$$

Hence we get (22) up to a constant for the generalized case.

In the above course, we also use the technique of  $\delta$ -function and the derivative of generalized function. Mistakes occur again since  $f$  is not a good function. However, if  $|\mu| = 0$ , the above argument still holds as long as we require  $f$  to be a  $C^1$ -function.

We conclude that Wu and Schaback's error bound can be used only when  $|\mu| = 0$  with an additional restriction that both  $\psi$  and  $f$  should be in  $C^1$  classically. What's noteworthy is that one should avoid using this error bound when  $\psi$  is a thin-plate spline. As for multiquadrics and Gaussians, it still works.

The significance of derivatives should be explained here. We know that derivatives reflect smoothness. If the approximand and approximant can match in smoothness, it will be nice. I fear r.b.f. people probably overemphasize the infinitely differentiable multiquadrics. Of course, multiquadrics are a very useful tool in the multivariate scattered data interpolation. However, as pointed out by Yoon in [7], most data points do not arise from extremely smooth functions. If this is true, multiquadrics are not very satisfactory in most cases. An example can be given here. It's well known that Weierstrass raised a function which is everywhere continuous, but is everywhere undifferentiable. If we approximate such a function by multiquadrics, it's meaningless even though its error bound is small. Although this example is a bit too drastic, it provides an evidence that sometimes we do have to consider the matching of smoothness. What's noteworthy is that in the practical problems of differential equations, most solutions are at most  $C^2$  or  $C^3$  functions. They are not very smooth.

I personally believe that unsmooth functions should be approximated by unsmooth functions. In fact, this is the central idea of this paper.

## 4 Comparison of the function spaces

In Wu and Schaback's approach, the space of approximated functions  $F_\phi$  is defined to be the set of all functions from  $R^n$  into  $R$  which has a generalized Fourier transform  $\hat{f}$  coinciding on  $R^n - \{0\}$  with a continuous function satisfying

$$\int_{R^n} |\hat{f}(t)|^2 (\hat{\psi}(t))^{-1} dt < \infty$$

for  $\psi(x) = \phi(\|x\|)$  satisfying (15) and (16).

Many people mistake Wu and Schaback's space for Madych and Nelson's space of approximands. This may lead to a disaster.

Madych and Nelson's space [4],[5] is much more complicated to describe than that of Wu and Schaback. However, L-T. Luh has made a clear characterization for Madych and Nelson's space which is easier to understand. Let's follow Luh's notation and definitions in [2] and [3].

In the following definitions,  $\Omega$  denotes an arbitrary subset of  $R^n$  and  $P_m^n$  denotes the set of  $n$ -variable polynomials of order not exceeding  $m$ .

**Definition 4.1**  $(P_m^n)_\Omega^\perp := \{\sum_{i=1}^N c_i \delta_{x_i} | c_i \in C, x_i \in \Omega \text{ for } 1 \leq i \leq N, \sum_{i=1}^N c_i p(x_i) = 0 \text{ for all } p \in P_m^n\}$  where  $\delta_{x_i}$  is a measure defined by  $\delta_{x_i}(E) = \begin{cases} 1 & \text{if } x_i \in E \\ 0 & \text{if } x_i \notin E \end{cases}$  for any subset  $E$  of  $R^n$ .

We can construct a norm in  $(P_m^n)_\Omega^\perp$ . If we define

$$(\lambda, \mu)_\psi := \sum_{i=1}^N \sum_{j=1}^M \lambda_i \bar{\mu}_j \psi(x_i - y_j)$$

for  $\lambda, \mu \in (P_m^n)_\Omega^\perp$  with  $\lambda = \sum_{i=1}^N \lambda_i \delta_{x_i}$ ,  $\mu = \sum_{j=1}^M \mu_j \delta_{y_j}$ , it's easily seen  $(\cdot, \cdot)_\psi$  is an inner product and induces a norm in  $(P_m^n)_\Omega^\perp$  as long as  $\psi$  is c.p.d. of order  $m$  in  $\Omega$ .

**Definition 4.2**  $\overline{(P_m^n)_\Omega^\perp}$  is the Hilbert space completion of  $(P_m^n)_\Omega^\perp$  under the norm induced by  $\psi$ .

**Definition 4.3**  $F_{\phi, \Omega} := \{f(x) = (\mu, \delta_{(x)})_\psi | \mu \in \overline{(P_m^n)_\Omega^\perp}, x \in \Omega, (\cdot, \cdot)_\psi \text{ is the inner product in } \overline{(P_m^n)_\Omega^\perp} \text{ induced by } \psi\}$  where  $\delta_{(x)} := \delta_x - \sum_{i=1}^{m'} l_i(x) \delta_{x_i}$  with  $l_1, \dots, l_{m'}$  a basis of  $P_m^n$  and  $x_1, \dots, x_{m'} \in \Omega$  satisfying  $l_i(x_j) = \delta_{ij}$ .

**Definition 4.4** Let  $\psi$  be c.p.d. of order  $m \geq 0$  in  $\Omega$ . Then

$$C_\psi(\Omega) = P_m^n(\Omega) \oplus \overline{F_{\psi, \Omega}}$$

where  $\oplus$  denotes the direct sum of two linear spaces.

In fact, Madych and Nelson's space of approximands is just  $C_\psi(R^n)$ .

In [6] Wu and Schaback say that the completion of  $F_\phi$  is just  $\overline{F_{\psi, R^d}}$ . Some authors even say that  $F_\phi$  is equal to  $C_\psi(R^d)$ . See [7] and [8].

In order to make a lucid clarification, I must quote two theorems of [4].

**Theorem 4.1** Let  $\psi$  be continuous and conditionally positive definite of order  $m$ . Then it is possible to choose a positive Borel measure  $\mu$  on  $R^n - \{0\}$ , constants  $a_\gamma$ ,  $|\gamma| \leq 2m$  and a function  $\chi$  in  $\mathcal{D}$  such that:  $1 - \hat{\chi}(\xi)$  has a zero of order  $2m + 1$  at  $\xi = 0$ ; both of the integrals  $\int_{0 < |\xi| < 1} |\xi|^{2m} d\mu(\xi)$ ,  $\int_{|\xi| \geq 1} d\mu(\xi)$  are finite; for all  $\sigma \in \mathcal{D}$ ,

$$\int \psi(x) \sigma(x) dx = \int [\hat{\sigma}(\xi) - \hat{\chi}(\xi) \sum_{|\gamma| < 2m} D^\gamma \hat{\sigma}(0) \frac{\xi^\gamma}{\gamma!}] d\mu(\xi) + \sum_{|\gamma| \leq 2m} D^\gamma \hat{\sigma}(0) \frac{a_\gamma}{\gamma!}. \quad (26)$$

This uniquely determines the measure  $\mu$  and the constants  $a_\gamma$  for  $|\gamma| = 2m$ . In addition, for every choice of complex numbers  $c_\alpha$ ,  $|\alpha| = m$ ,

$$\sum_{|\alpha|=m} \sum_{|\beta|=m} a_{\alpha+\beta} c_\alpha \bar{c}_\beta \geq 0.$$

**Theorem 4.2** *Let  $m, \psi, \mu$  and  $a_\gamma$  be as in Theorem 4.1. Assume  $d\mu(\xi) = w(\xi)d\xi$  and  $a_\gamma = 0$  for all  $|\gamma| = 2m$ . Let  $\rho$  be the Borel measure on  $R^n$  defined by  $d\rho(\xi) = r(\xi)d\xi$ , where  $r$  is defined by*

$$r(\xi) = \frac{1}{(2\pi)^{2n}|\xi|^{2m}w(-\xi)},$$

*with  $r(\xi) = \infty$  when  $w(-\xi) = 0$ . Then  $f \in C_\psi(R^n)$  if and only if  $f \in \mathcal{S}'$  and  $(D^\alpha f)^\wedge \in L^2(\rho)$  for every  $|\alpha| = m$ . In that case,  $(f, f)_\psi$  is given by*

$$\sum_{|\alpha|=m} \frac{m!}{\alpha!} \|(D^\alpha f)^\wedge\|_{L^2(\rho)}^2 = \int |g|^2 d\mu = (f, f)_\psi$$

*where  $g$  is defined in Proposition 3.2 of [5].*

By Theorem 4.2, one easily finds that Wu and Schaback's  $F_\phi$  is equal to a subset of Madych and Nelson's  $C_\psi(R^n)$  when  $m = 0$ , with the same norm. Here, one should be careful. Wu and Schaback's space consists only of real-valued functions. However, Madych and Nelson's space contains complex-valued functions. When  $m > 0$ , the two spaces cannot be compared. For  $m > 0$ , the Hilbert space completion of  $F_\phi$  still has nothing to do with  $C_\psi(R^n)$ .

Now, another big problem of Wu and Schaback's error bound can be pointed out. That is, there is no reason to believe that the generalized derivative  $f^{(\mu)}(x)$  in (23) coincides with a nicely behaving classical function. This question is absolutely nontrivial. Although some people may add this condition as an additional requirement, this will become a drawback because it's rather restrictive. Madych and Nelson deal with this problem perfectly in their paper [5]. In their theory the generalized derivative coincides naturally with a nicely behaving classical function. However, it's suitable only for  $C_\psi(R^n)$ . Moreover, Madych and Nelson's space  $C_\psi(R^n)$  seems to be more meaningful than Wu and Schaback's space  $F_\phi$  for two reasons. First, it has a natural interpretation in physics. Second, it contains a dense subset which consists of natural extensions of the univariate cubic spline. The two reasons are very strong. All these can be found in [4].

In fact, not only Wu, Schaback and Yoon, but also many others misunderstand the two spaces. The main reason is that Madych and Nelson's papers are very difficult to read. In order to get a better understanding to their papers, it might be helpful to read [2] and [3] first.

Finally, I would like to point out a key point of the comparison of the two spaces and the two error bounds. It's well known that there are two notions of generalized Fourier transform. One is the Russian notion of Gelfand and Schilov. The other is the French notion of Schwartz. Many people think the notion used by Madych and Nelson in [5] is the Russian notion. In fact, this is wrong. What they use is still the French notion. This can be seen in page 213 and page 219 of [5]. Madych and Nelson studied Duchon's works and tried to promote it. That

probably is the reason why they use the French notion of generalized Fourier transform. Unfortunately, in the bibliography of [5], only Gelfand and Vilenkin's book "Generalized Functions", Vol.4, 1964 appears. None of Schwartz, Jones, or Stein/Weiss's book appears. This misleads people into thinking that the notion used by them is the Gelfand/Schilov's definition. This seems to be part of the reasons many people do not like Madych and Nelson's approach. They prefer Wu and Schaback's approach which uses the better digestible French notion.

## References

- [1] Jones, D.S., *The Theory of Generalized Functions*, Cambridge University Press, Cambridge 1982
- [2] Luh, Lin-Tian, *The equivalence theory of native spaces*, Approx. Theory Appl., 17:1,2001, pp. 76-96.
- [3] Luh, Lin-Tian, *The embedding theory of native spaces*, Approx. Theory Appl., 17:4, 2001, pp. 90-104.
- [4] Madych, W.R., and Nelson, S.A., *Multivariate interpolation and conditionally positive definite function I*, Approx. Theory Appl.,4(1988), pp. 77-89.
- [5] Madych, W.R., and Nelson, S.A., *Multivariate interpolation and conditionally positive definite function II*, Math. Comp.,54(1990),pp. 211-230.
- [6] Wu, Z., and Schaback, R., *Local error estimates for radial basis function interpolation of scattered data*, IMA J. Numer. Anal. 13(1993),13-27.
- [7] Yoon, J., *Spectral approximation orders of radial basis function interpolation on the Sobolev space*, SIAM J. MATH. ANAL. Vol.33,No.4,2001,pp.946-958.
- [8] Yoon, J., *Interpolation by radial basis functions on Sobolev space*, To appear in J. of Approx. Theory, 2001.

Department of Applied Mathematics,  
 Providence University,  
 Shalu Town,  
 Taichung County,  
 Taiwan

Email:ltluh@pu.edu.tw